

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
ĐẠI HỌC ĐÀ NẴNG**

CAO BÁ THÀNH

**NGHIÊN CỨU ỨNG DỤNG ĐƯỜNG CONG THAM SỐ
B-SPLINE VÀO NHẬN DẠNG CHỮ SỐ VIẾT TAY**

Chuyên ngành: KHOA HỌC MÁY TÍNH

Mã ngành: 60.48.01

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT

Chương trình được hoàn thành tại

ĐẠI HỌC ĐÀ NẴNG

Người hướng dẫn khoa học: **TS. Nguyễn Tấn Khôi**

Phản biện 1: **PGS.TS. Tăng Tấn Chiến**

Phản biện 2: **TS. Trương Công Tuấn**

Luận văn được bảo vệ trước Hội đồng chấm Luận văn tốt nghiệp thạc sĩ Kỹ thuật họp tại Đại học Đà Nẵng vào ngày 03 tháng 03 năm 2012

MỞ ĐẦU

1. Lý do chọn đề tài

Hiện nay công nghệ thông tin phát triển rất mạnh, đã đi sâu vào mọi thành phần trong xã hội và máy tính đã trở thành công cụ đắc lực không thể thiếu trong nhiều lĩnh vực, nó giúp chúng ta lưu trữ các dữ liệu một cách hiệu quả và đặc biệt là các tài liệu đã được số hóa. Trong nhiều lĩnh vực thuộc ngành Công nghệ thông tin thì Trí tuệ nhân tạo là lĩnh vực được quan tâm hàng đầu hiện nay và được đánh giá sẽ phát triển rất mạnh trong tương lai. Lĩnh vực Trí tuệ nhân tạo nói chung, bài toán nhận dạng chữ viết tay nói riêng đã được nhiều nhà nghiên cứu quan tâm vì ý nghĩa thực tiễn to lớn của nó.

Trong điều kiện thực tế hiện nay còn rất nhiều các trường hợp nhận thông tin quản lý thông qua các biểu mẫu và ngày một nhiều các tài liệu quan trọng vẫn còn được lưu trữ trên giấy. Trong các trường học phiếu điểm hay đăng ký thông tin đầu vào của học sinh, sinh viên được lưu trữ ở các biểu mẫu trên giấy, cũng như các cơ quan quản lý khi nhận thông tin của người dân để xử lý cũng từ các biểu mẫu ... Để phục vụ cho công tác chỉnh sửa, tra cứu, in ấn và quản lý tốt thì các cơ quan, trường học... cần phải số hóa các thông tin này để lưu trữ trên máy tính.

Để giải quyết bài toán trên, thông thường có hai cách thực hiện:

Cách 1: Quét tài liệu viết tay trên giấy thành file ảnh và lưu vào máy tính. Nhưng như thế thì không thể chỉnh sửa, tra cứu hay quản lý tốt các thông tin này.

Cách 2: Số hóa thủ công các tài liệu viết tay trên giấy. Cách này mất rất nhiều thời gian và chi phí.

Từ kết quả trên, chúng ta thấy rằng số hóa tài liệu cần phải có công cụ tự động là cấp bách, nhằm giảm thiểu thời gian và chi

phí. Phương pháp B-Spline tỏ ra ưu thế hơn một số phương pháp khác ở chỗ, khi ra quyết định nhận dạng của một số phương pháp khác được cài đặt tinh trong chương trình, muốn bổ sung thêm một số mẫu mới phải thiết kế lại chương trình. Trong khi đó với phương pháp B-Spline chỉ cần ta cung cấp các mẫu mới và cho máy học là ta bổ sung vào bộ nhớ những kiểu dữ liệu mới, việc này không ảnh hưởng đến cấu trúc chương trình ban đầu. Xuất phát từ những vấn đề trên, việc tiến hành đề tài “Nghiên cứu ứng dụng đường cong tham số B-Spline vào nhận dạng chữ số viết tay” là rất cần thiết.

2. Mục tiêu nhiệm vụ của đề tài

Mục tiêu của đề tài là nghiên cứu ứng dụng đường cong tham số B-spline để máy học các ký tự mẫu có cấu trúc mở và xây dựng hệ thống cho phép nhận dạng một số tài liệu từ file ảnh nhằm giúp số hóa tài liệu tự động và nhanh chóng.

Xây dựng ứng dụng cho phép đưa vào tập tin bitmap của một bảng điểm và xuất ra mã số học sinh, điểm tương ứng của học sinh đó ở trường THPT số 1 Nghĩa Hành, từ đó triển khai cho các trường.

Các nhiệm vụ cụ thể của đề tài bao gồm

Nghiên cứu lý thuyết về đường cong tham số B-spline.

Nghiên cứu lý thuyết xử lý ảnh, áp dụng vào bước tiền xử lý ảnh (ảnh Bitmap của ký tự).

Nghiên cứu cho máy học và nhận dạng ký tự bằng phương pháp sử dụng đường cong tham số B-spline qua các tham số đặc trưng như: đỉnh điều khiển, vector nút.

Nghiên cứu xây dựng ứng dụng nhận dạng chữ và số viết tay sử dụng đường cong B-spline và ứng dụng vào đề nhập điểm tự động.

3. Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu

Đường cong tham số B-spline.

Lý thuyết xử lý ảnh.

Kỹ thuật nhận dạng ký tự viết tay bao gồm: nhận dạng chữ và số viết tay.

Ngôn ngữ lập trình C#.

Phạm vi nghiên cứu

Nghiên cứu các lý thuyết toán học của đường cong B-Spline có thể áp dụng vào phương pháp nhận dạng ký tự viết tay.

Nghiên cứu các phương pháp, thuật toán đơn giản phục vụ cho việc tiền xử lý ảnh ký tự.

Cho phép người dùng đưa ảnh ký tự để máy học cũng như nhận dạng.

Đối với nhận dạng chữ viết tay chỉ cho máy học và nhận dạng các ký tự số từ “0” đến “9” và các ký tự chữ cái.

4. Phương pháp nghiên cứu

Đề tài nghiên cứu về đường cong B-spline và xử lý ảnh để xử lý dữ liệu ảnh đầu vào. Từ đó đề xuất phương pháp, giải thuật và xây dựng ứng dụng nhận dạng ký tự viết tay. Để làm rõ vấn đề trên, luận văn sử dụng phương pháp nghiên cứu lý thuyết và thực nghiệm.

Nghiên cứu lý thuyết

Đường cong tham số B-spline.

Xử lý ảnh.

Các mô hình tổ chức lưu trữ dữ liệu.

Nghiên cứu thực nghiệm

Thu thập tài liệu và thông tin liên quan đến đề tài.

Lựa chọn mô hình, cách tiếp cận phù hợp với nội dung.

Đề xuất qui trình nhận dạng dùng đường cong tham số B-spline.

Triển khai xây dựng chương trình.

Kiểm thử, nhận xét và đánh giá kết quả của hệ thống.

5. Ý nghĩa khoa học và thực tiễn

Áp dụng kỹ thuật đường cong tham số B-spline vào nhận dạng chữ viết tay và ứng dụng cho nhận dạng điểm số từ bảng điểm ở trường học.

Ứng dụng nhận dạng bảng điểm phục vụ nhập điểm tự động tại trường THPT số 1 Nghĩa Hành.

Trợ giúp cho giáo viên và nhà trường đỡ mất thời gian và kinh phí cho việc nhập các bảng điểm vào máy tính.

6. Bố cục của luận văn

Bố cục luận văn gồm ba chương được tóm tắt như sau

Chương 1 - Cơ sở lý thuyết

Chương này trình bày những cơ sở lý thuyết có liên quan đề tài như: Tìm hiểu đường cong B-spline và xử lý ảnh để áp dụng cho nhận dạng chữ viết tay.

Chương 2 - Giải pháp ứng dụng đường cong B-spline vào nhận dạng

Trong chương này, giới thiệu giải pháp quá trình nhận dạng chữ viết tay, phương pháp và giải thuật dùng trong nhận dạng chữ viết.

Chương 3 - Xây dựng hệ thống và thử nghiệm

Phân tích các chức năng của hệ thống, thiết kế kiến trúc hệ thống và thực hiện xây dựng ứng dụng theo cách thức áp dụng đường cong B-spline vào nhận dạng chữ viết tay, sau đó thử nghiệm và đánh giá kết quả đạt được của chương trình.

CHƯƠNG 1 - CƠ SỞ LÝ THUYẾT

Chương này giới thiệu về cơ sở lý thuyết có liên quan đến đề tài, làm nền tảng nhằm hỗ trợ xây dựng ứng dụng đặt ra của đề tài sau này, gồm các nội dung sau

- Giới thiệu một số phương pháp nhận dạng chữ viết tay đang được sử dụng hiện nay.

- Cơ sở lý thuyết về đường cong tham số B-spline.

- Giới thiệu cách thu nhận ảnh, tách ký tự, nhị phân hóa ký tự, tìm xương và giảm số điểm biểu diễn ký tự.

1.1. Một số phương pháp nhận dạng chữ viết tay

1.1.1. Phương pháp nhận dạng sử dụng lưới

1.1.2. Phương pháp nhận dạng bằng chia miền và đo mật độ

1.1.3. Phương pháp nhận dạng bằng phân tích đường đơn

1.1.4. Phương pháp nhận dạng bằng phân tích đường biên

1.2. Xử lý ảnh cho nhận dạng chữ viết

1.2.1. Những vấn đề cơ bản trong xử lý ảnh

1.2.1.1. Điểm ảnh

1.2.1.2. Độ phân giải của ảnh

1.2.1.3. Mức xám của ảnh

1.2.1.4. Định nghĩa ảnh số

1.2.1.5. Quan hệ giữa các điểm ảnh

1.2.2. Lọc nhiễu

Để làm được điều này chúng tôi dùng một bộ lọc giống như lọc Medial nghĩa là dùng một ma trận 3x3 quét khắp ảnh bitmap. Thực ra bộ lọc này chính là ánh xạ đi từ ảnh nguồn A sang file ảnh đích B định bởi [1]

$$W(i, j) = \begin{cases} 1 & \text{nếu } \sum_{i=-1}^2 \sum_{j=-1}^2 W(i, j) \geq 3 \\ 0 & \text{trường hợp khác} \end{cases}$$

Thế nhưng việc lọc nhiễu trong fiel ảnh có thể làm mất dấu chấm của chữ i hoặc chữ j. Do đó trong một số trường hợp không thể khử nhiễu[1][3].

1.2.3. Nhị phân hóa ký tự

Trong đồ họa máy tính, màu của một điểm ảnh với định dạng ảnh RGB được xác định bởi ba chỉ số: R: Red (đỏ), G: Green (Lục - xanh lá cây), B: Blue (Lam - xanh da trời). R, G, B nằm trong khoảng [0..255].

Để xác định độ sáng của điểm ảnh, ta sử dụng công thức[1]

$$L = 0.299 * R + 0.587 * G + 0.114 * B$$

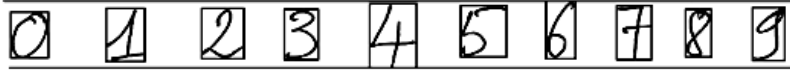
Ta sẽ quy ước nếu $L \geq 0.8$ sẽ là pixel trắng và ngược lại sẽ là pixel đen, từ đó ta có ma trận nhị phân của hình ảnh.

1.2.4. Phân tách ký tự

Quá trình phân tách ký tự, trải qua hai bước: Xác định tọa độ của các hàng ký tự và xác định tọa độ cực trái và cực phải của các ký tự trên một hàng.

Để xác định tọa độ trên và dưới của một hàng ký tự, ta sử dụng một đường thẳng quét từ trên xuống dưới ma trận nhị phân của ký tự. Nếu đường thẳng này gặp một pixel đen (0) ta đánh dấu tọa độ cực trên của hàng. Ta tiếp tục quét cho đến khi gặp một hàng trắng (không có pixel đen nào trên hàng), ta đánh dấu cực dưới của hàng, từ đó ta xác định được tọa độ của hàng, tiếp tục quá trình cho đến hết chiều dài thẳng đứng của ma trận nhị phân. Áp dụng phương pháp tương tự theo chiều dọc cho từng hàng ta xác định được tọa độ cực

trái và cực phải của từng ký tự. Xác định các cực của ký tự được minh hoạ hình 1.5 sau



Hình 1.5 Phân tách ký tự

1.2.5. Tìm xương ký tự

1.2.6. Giảm số điểm biểu diễn ký tự

1.3. Đường cong B-spline

1.3.1. Phương trình và tính chất của đường cong B-spline

Đường cong B-Spline xác định bởi công thức [10][11][12]

$$P(t) = \sum_{i=0}^n N_{i,k}(t) B_i \quad t_{\min} \leq t < t_{\max}, \quad 2 \leq k \leq n+1 \quad (1.1)$$

Trong đó:

- B_i là tập hợp $n+1$ điểm kiểm soát, k là cấp của đường cong
- $N_{i,k}$ là hàm cơ sở B-Spline bậc $k-1$

Hàm cơ sở B-Spline thứ i là $N_{i,k}(t)$ được định nghĩa bởi phương trình đệ quy sau

$$N_{i,0}(t) = \begin{cases} 1 & \text{nếu } t_i \leq t < t_{i+1} \\ 0 & \text{trường hợp khác} \end{cases}$$

$$\text{với } N_{i,k}(t) = \frac{(t - t_i)}{t_{i+k-1} - t_i} N_{i,k-1}(t) + \frac{(t_{i+k} - t)}{t_{i+k} - t_{i+1}} N_{i+1,k-1}(t) \quad (1.2)$$

(quy ước: $0/0 = 0$)

vector nút là $[t_0, \dots, t_m]$ là một dãy số nguyên tăng dần $t_0 < t_1 < \dots < t_{m-1}$

1.3.2. Các dạng vector nút

1.3.2.1. Vector nút dạng đồng nhất

1.3.2.2. Vector nút dạng đồng nhất mở

1.3.3. Điều khiển hình dạng của đường cong B-Spline

1.3.4. Thêm nút vào vector nút

Cho đường cong B-Spline được định nghĩa bởi công thức[10][11][12]

$$P(t) = \sum_{i=0}^n N_{i,k}(t)B_i$$

Với vector nút $[T] = [t_0 \ t_1 \ \dots \ t_{n+k}]$

Sau khi thêm nút vào vector nút, ta có một đường cong mới

$$R(y) = \sum_{j=0}^m M_{j,k}(y)c_j$$

Với vector nút $[Y] = [y_0 \ y_1 \ \dots \ y_{m+k}]$

Giả sử $m > n$. Mục đích là xác định các đa giác kiểm soát mới C_j sao cho $P(t) = R(y)$. Theo thuật toán của Oslo[12] ta có

$$C_j = \sum_{i=0}^n \alpha_{i,j}^k B_i \quad 0 \leq i < n, 0 \leq j < m \quad (1.3)$$

Hàm $\alpha_{i,j}^k$ được tính như sau

$$\alpha_{i,j}^k = \begin{cases} 1 & \text{nếu } t_i \leq y_j < t_{i+1} \\ 0 & \text{trường hợp khác} \end{cases}$$

$$\text{và } \alpha_{i,j}^k = \frac{(y_{j+k-1} - t_i)}{t_{i+k-1} - t_i} \alpha_{i,j}^{k-1} + \frac{(y_{j+k-1} - t_{i+k})}{t_{i+k} - t_{i+1}} \alpha_{i+1,j}^{k-1} \quad (1.4)$$

$$\sum_{i=0}^n \alpha_{i,j}^k = 1$$

1.3.5. Các phép biến đổi affine đối với đường cong B-spline

1.3.5.1. Phép quay quanh gốc tọa độ

1.3.5.2. Phép lấy đối xứng qua trục tọa độ

1.3.5.3. Phép biến đổi tỷ lệ hay phóng to, thu nhỏ (Scale)

1.3.5.4. Phép biến đổi tuyến tính và tọa độ Homogeneous

1.3.5.5. Phép quay quanh một điểm bất kỳ M

1.3.5.6. Phép lấy đối xứng qua đường thẳng bất kỳ

1.4. Kết chương

Đường cong B-spline có nhiều ưu điểm hơn so với các đường cong khác. Thông qua các đa thức tham số xác định riêng rẽ trên một số điểm kiểm soát lân cận với số bậc tùy ý không phụ thuộc vào số lượng các điểm kiểm soát, có nghĩa khi dịch chuyển điểm kiểm soát của đường cong thì chỉ một vài phân đoạn lân cận của điểm kiểm soát đó bị ảnh hưởng chứ không phải toàn bộ đường cong.

Đường cong B-spline bất biến đối với các phép biến đổi affin và đặc biệt là bất biến với phép biến đổi homogeneous (tức là phép chiếu phối cảnh) nghĩa là khi ta muốn biến đổi đường cong B-spline ta chỉ cần thực hiện phép biến đổi trên các điểm kiểm soát của đường cong đó.

CHƯƠNG 2 - ỨNG DỤNG ĐƯỜNG CONG B-SPLINE VÀO NHẬN DẠNG CHỮ VIẾT TAY

Trong chương này, tập trung giải quyết các công việc sau

- Đưa ra ý tưởng và giải pháp nhận dạng chữ viết tay ứng dụng đường cong B-spline.

- Đưa ra giải thuật và phương pháp sử dụng trong quá trình nhận dạng chữ viết tay.

2.1. Giải pháp nhận dạng

2.1.1. Ý tưởng nhận dạng

Trong quá trình nghiên cứu đường cong B-spline chúng tôi nhận ra rằng phương pháp xây dựng đường cong là khá hiệu quả. Theo đó ta chỉ cần xác định một số hữu hạn các điểm kiểm soát (Control points) và từ các điểm kiểm soát này ta dễ dàng xây dựng nên đường cong B-spline.

Tính chất quan trọng của đường cong B-spline là bất biến với các phép biến đổi Affine (phép tịnh tiến, phép co và phép quay) nghĩa là chúng ta muốn co đường cong B-spline ta chỉ cần co các điểm kiểm soát của chúng mà đường cong không bị biến dạng.

Ý tưởng chính

- Cho máy học các ký tự mẫu: đọc từ tập tin ảnh các ký tự cần học sau đó xử lý, xác định các vector kiểm soát của chúng và lưu lên tập tin dữ liệu.

- Đọc ký tự muốn nhận dạng từ tập tin ảnh ký tự, xử lý và tìm các vector kiểm soát của chúng. Đối sánh các vector kiểm soát của ký tự cần nhận dạng với các vector kiểm soát của ký tự mà máy đã học và đưa ra kết quả.

Một số trường hợp cần phải xét thêm là các ký tự có thể không bằng nhau về kích thước, có thể bị viết nghiêng...

Trường hợp 1: Ký tự có kích thước khác nhau

Dựa trên tính chất bất biến với phép biến đổi Affine (phép tịnh tiến, phép thu nhỏ, phép phóng to, phép quay). Ta chỉ cần co hay phóng lớn ký tự đó về một khung chuẩn, nghĩa là ta chỉ việc phóng to hay thu nhỏ các điểm kiểm soát của chúng.

Trường hợp 2: Ký tự không chồng khít lên nhau ở khung chuẩn

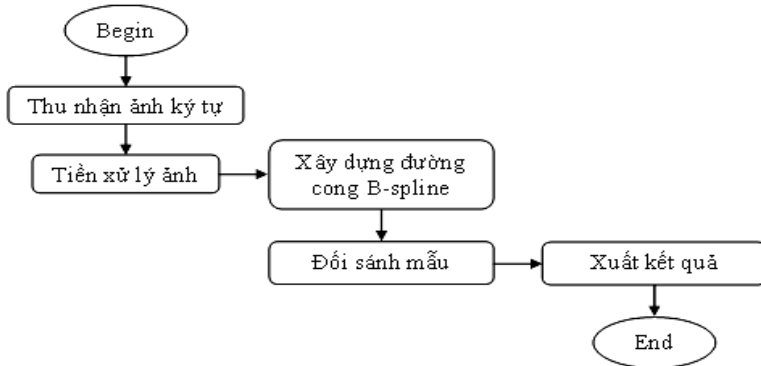
Lúc này chúng ta có thể tịnh tiến thăm dò (tịnh tiến qua trái, qua phải, lên hay xuống) cho đến khi chúng chồng khít lên nhau, lúc này sai số khoảng cách là nhỏ nhất.

Trường hợp 3: Ký tự viết nghiêng

Trong trường hợp này ta có thể sử dụng phép quay ký tự thăm dò. Nếu ta quay theo chiều dương mà sai số khoảng cách (dùng độ đo Euclide) so với ký tự mẫu mà tốt hơn thì ta tiếp tục quay theo hướng này, không tốt hơn thì ta quay theo chiều ngược lại.

2.1.2. Phương pháp nhận dạng ký tự viết tay

Phương pháp ứng dụng đường cong B-spline vào nhận dạng chữ viết tay bao gồm các bước được mô tả như sau



Hình 2.7 Sơ đồ nhận dạng ký tự viết tay sử dụng đường cong B-spline

2.1.2.1. Thu nhận ảnh ký tự

Ảnh ký tự có thể được thu nhận bằng máy Scanner, Webcam, hoặc các thiết bị thu nhận ảnh thông dụng khác và được lưu dưới dạng file ảnh.

2.1.2.2. Tiền xử lý ảnh

Quá trình tiền xử lý ảnh ký tự bao gồm các bước sau:

- Lọc nhiễu là công đoạn quan trọng trong xử lý ảnh, làm cho ảnh tốt hơn để ứng dụng vào xác định số hàng và số cột đạt hiệu quả cao hơn.

- Nhị phân ảnh là chuyển ảnh bitmap thành ma trận nhị phân để ứng dụng vào các công đoạn sau hiệu quả hơn, như tìm xương, phân tách ký tự...

- Phân tách thông tin ra khỏi đối tượng hình học nhằm mục đích tách các ký tự ra khỏi đối tượng hình học để lấy riêng phần ký tự nằm trong các khung biểu mẫu.

- Phân tách ký tự là tách từng ký tự một để áp dụng vào tìm xương và tìm các vectơ kiểm soát sau này.

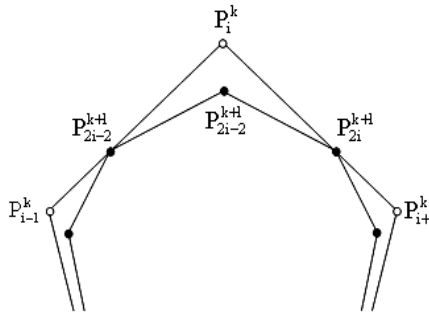
- Tìm xương nhằm mục đích giảm thiểu số lượng điểm biểu diễn của đường cong, để thuận tiện cho việc tìm điểm kiểm soát của ký tự và đối sánh sau này.

- Giảm số lượng điểm biểu diễn ảnh mục đích là giảm thiểu không quan lưu trữ và thuận tiện cho việc đối sách trong nhận dạng sau này.

2.1.2.3. Xây dựng đường cong B-Spline

Có rất nhiều cách để xây dựng đường cong B-spline dựa trên tập điểm kiểm soát của chúng, trong luận văn đề xuất áp dụng phương pháp phân rã (subdivision) để xây dựng đường cong tham số B-spline[15].

Phương pháp phân rã đường cong B-spline bậc 3 dựa trên thuật toán phân chia khung đa giác điều khiển. Bằng cách chèn liên tiếp các đỉnh mới vào khung đa giác ban đầu ta được khung đa giác mới có nhiều điểm hơn, và tiệm cận đường cong B-spline[15].



Hình 2.8 Xây dựng đường cong dựa trên tập các điểm kiểm soát

- Thuật toán phân chia (Subdivision) khung đa giác điều khiển[15]

Cho đa giác có n điểm kiểm soát $P^k = \{P_0^k, P_1^k, \dots, P_{n-1}^k\}$ ($k \in \mathbb{N}$).
 Các bước xây dựng đường cong B-spline bậc ba thực hiện như sau

Xác định các điểm mới giữa mỗi cạnh của đa giác kiểm soát dựa trên công thức [2.2a và 2.2b] sau

Ở bước $(k + 1)$ chúng ta có được đường cong $P^{k+1} = \{P_i^k\}_{i=0..m}$ và điểm $\{P_i^k\}$ được tính như sau

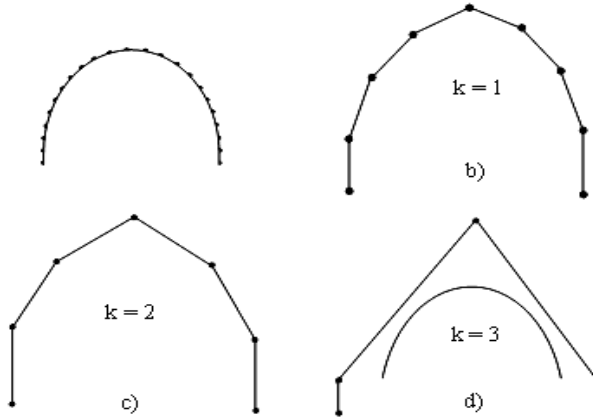
$$P_{2i-1}^{k+1} = \frac{1}{8}P_{i-1}^k + \frac{3}{4}P_i^k + \frac{1}{8}P_{i+1}^k \quad \text{cho } i = 1 \dots n-2 \quad (2.2a)$$

$$P_{2i}^{k+1} = \frac{1}{2}P_i^k + \frac{1}{2}P_{i+1}^k \quad \text{cho } i = 0 \dots n-2 \quad (2.2b)$$

Khi số bước k tăng lên, các điểm kiểm soát sẽ tiệm cận với đường cong B-spline.

Ở trên chúng ta đã tạo được đường cong tham số B-spline áp dụng phương pháp phân rã. Bài toán ngược lại, tức là từ tập các điểm nằm trên đường cong, cần phải xác định các điểm kiểm soát của chúng.

- Thuật toán tái tạo (Inverse Subdivision) khung điều khiển của đường cong B-spline từ một tập điểm cho trước[15]



Hình 2.9 Xác định điểm kiểm soát từ đường cong. a) các điểm trên đường cong. b) $k = 1$. c) $k = 2$. d) $k = 3$.

$k = 3$ ta thu được một khung điều khiển có 5 đỉnh, từ đó ta có thể xây dựng được đường cong B-spline bậc 3.

Từ các đỉnh ban đầu, tính ngược các đỉnh dựa trên công thức sau

$$P_i^k = -\frac{1}{2}P_{2i-2}^{k+1} + 2P_{2i-1}^{k+1} - \frac{1}{2}P_{2i+1}^{k+1} \quad \text{cho} \quad i = 1 \dots n-2 \quad (2.3a)$$

$$P_i^k = 2P_{2i}^{k+1} - P_{i+1}^k \quad \text{cho} \quad i = 0 \quad (2.3b)$$

$$P_i^k = 2P_{2i-2}^{k+1} - P_{i-1}^k \quad \text{cho} \quad i = n-1 \quad (2.3c)$$

2.1.2.4. Đối sánh ký tự cần nhận dạng với mẫu cho trước

Quy trình đối sánh mẫu bao gồm các bước sau[11][12]

- Dùng độ đo khoảng cách Euclide để đối sánh với các vector kiểm soát của ký tự cần nhận dạng với tập vector mẫu để tìm ra ký tự giống nhất.

- Trường hợp các vector của ký tự cần nhận dạng ít hơn các vector mẫu thì làm mịn đường cong bằng cách thêm điểm kiểm soát.

2.1.2.5. Xuất kết quả

2.2. Phương pháp và giải thuật sử dụng trong nhận dạng

2.2.1. Phương pháp phân tách ký tự ra khỏi đối tượng hình học

Xác định các đường ngang của bảng điểm

1) Sử dụng một đường thẳng ngang quét từ trên xuống đến khi đụng đường kẻ ngang của bảng điểm là liên tiếp các pixel đen (0) theo chiều ngang của đường thẳng.

2) Tiếp tục quét đến khi nét của đường thẳng gặp liên tiếp các pixel trắng (1) thì đánh dấu tọa độ cực trên của đối tượng hình học.

3) Tiếp tục cho đường thẳng quét đến khi gặp liên tiếp các pixel đen (0) thì đánh dấu tọa độ cực dưới của đối tượng hình học.

4) Lặp lại bước 1 cho đến hết chiều dài đường thẳng đứng.

Tương tự phương pháp trên ta xác định được cực trái và cực phải của đối tượng hình học.

2.2.2. Thuật toán phân tách ký tự

Để phân tách ký tự, ta phải trải qua hai giai đoạn: Xác định tọa độ của các hàng ký tự và xác định tọa độ cực trái và cực phải của các ký tự trên một hàng.

Xác định tọa độ cực trên và cực dưới của một hàng ký tự

1) Ta sử dụng một đường thẳng ngang quét từ trên xuống dưới ma trận nhị phân của ký tự.

2) Nếu đường thẳng này gặp một pixel đen (0) ta đánh dấu tọa độ cực trên của hàng.

3) Nếu đường thẳng này gặp một hàng trắng (không có pixel đen nào trên hàng), ta đánh dấu cực dưới của hàng.

4) Lặp lại bước 1 đến hết chiều dài đường thẳng đứng của ma trận nhị phân ảnh.

Tương tự phương pháp trên ta xác định được cực trái và cực phải của ký tự



Hình 2.12 Phân tách ký tự

2.2.3. *Đo khoảng cách các vectơ kiểm soát*

Lần lượt đi qua các mẫu chuẩn đã được học, ta thực hiện co hoặc phóng to ký tự cần nhận dạng về khung chuẩn rồi mới bắt đầu đối sánh.

Sử dụng độ đo khoảng cách Euclide để đo sự sai khác giữa các vectơ kiểm soát với nhau. Gọi hai vectơ kiểm soát đó là $p = (p_1, p_2, \dots, p_n)$ và $q = (q_1, q_2, \dots, q_n)$, ta xem p và q là hai điểm trong không gian Euclide n chiều, thì khoảng cách từ p đến q được tính bằng [11][12]

$$E(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (2.4)$$

Trường hợp nếu độ dài của hai vectơ kiểm soát khác nhau, giả sử p có độ dài n , q có độ dài m với $m > n$ thì ta cần bổ sung vào p với $k = m - n$ điểm kiểm soát nữa để cho độ dài của hai vectơ bằng nhau. Cách thêm điểm kiểm soát tham khảo tại [11][12].

2.3. Kết chương

Trong chương này trình bày giải thuật sử dụng đường cong B-Spline vào nhận dạng chữ và số viết tay. Cho máy học các ký tự mẫu từ file bitmap, xử lý và xác định các vectơ kiểm soát của chúng và lưu lên file dữ liệu. Đọc ký tự muốn nhận dạng từ file bitmap, xử lý và tìm các vectơ kiểm soát của chúng, đối sánh các vectơ kiểm soát của ký tự cần nhận dạng với các vectơ kiểm soát của ký tự mà máy đã học và cho ra kết quả.

Phương pháp B-Spline tỏ ra ưu thế hơn một số phương pháp khác ở chỗ, khi ra quyết định nhận dạng của một số phương pháp

khác được cài đặt tinh trong chương trình, muốn bổ sung thêm một số mẫu mới phải thiết kế lại chương trình. Trong khi đó với phương pháp B-Spline chỉ cần ta cung cấp các mẫu mới và cho máy học là ta bổ sung vào bộ nhớ những kiểu dữ liệu mới, việc này không ảnh hưởng đến cấu trúc chương trình ban đầu.

CHƯƠNG 3 - XÂY DỰNG HỆ THỐNG NHẬN DẠNG CHỮ VIẾT TAY

Trong chương trước, luận văn đã giới thiệu về phương pháp nhận dạng ứng dụng đường cong tham số B-spline, xử lý dữ liệu từ file ảnh và một số giải thuật trong nhận dạng. Chương này tập trung vào việc phân tích và thiết kế hệ thống, xây dựng, thử nghiệm ứng dụng và từ đó đánh giá kết quả đạt được.

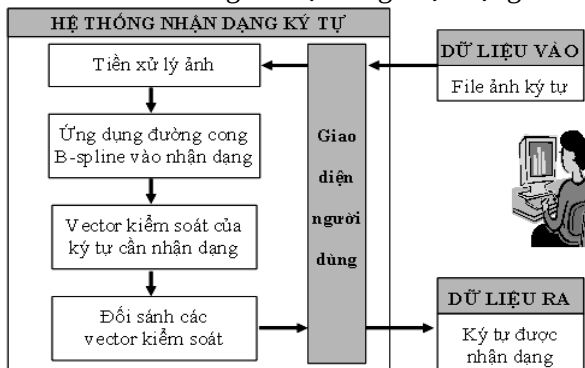
3.1. Phân tích và thiết kế hệ thống

3.1.1. Chức năng máy học ký tự mẫu

3.1.2. Chức năng nhận dạng

3.1.3. Kiến trúc tổng thể của hệ thống

Mô hình kiến trúc tổng thể hệ thống nhận dạng chữ viết tay.



Hình 3.3 Hệ thống nhận dạng ký tự.

3.2. Kịch bản sử dụng

3.3. Kết quả thực nghiệm

3.3.1. Nhận dạng chữ in

Cho máy học mẫu ký tự số in font Arial cỡ chữ 14.

Khi cho máy nhận dạng ký tự số in có cùng kích cỡ và font chữ thì máy nhận dạng đúng 100%.

Khi cho nhận dạng ký tự số in cùng font chữ nhưng kích cỡ 18 thì máy nhận dạng đúng 98%.

Khi cho nhận dạng ký tự số in cùng font chữ nhưng kích thước 10 thì máy nhận dạng đúng khoảng 90%.

3.3.2. Nhận dạng chữ viết tay

Cho máy học mẫu ký tự số viết tay từ "0" đến "9" mỗi ký tự viết 5 lần.

Mẫu máy học	Ký tự
1 1 1 1 1	1
2 2 2 2 2	2
3 3 3 3 3	3
4 4 4 4 4	4
5 5 5 5 5	5
6 6 6 6 6	6
7 7 7 7 7	7
8 8 8 8 8	8
9 9 9 9 9	9
0 0 0 0 0	0

Hình 3.7 Mẫu cho máy học ký tự số viết tay

Sau đó viết lại 10 ký tự số từ "0" đến "9" và cho nhận dạng thì máy nhận dạng cho kết quả như sau

0 1 2 3 4 5 6 7 8 9

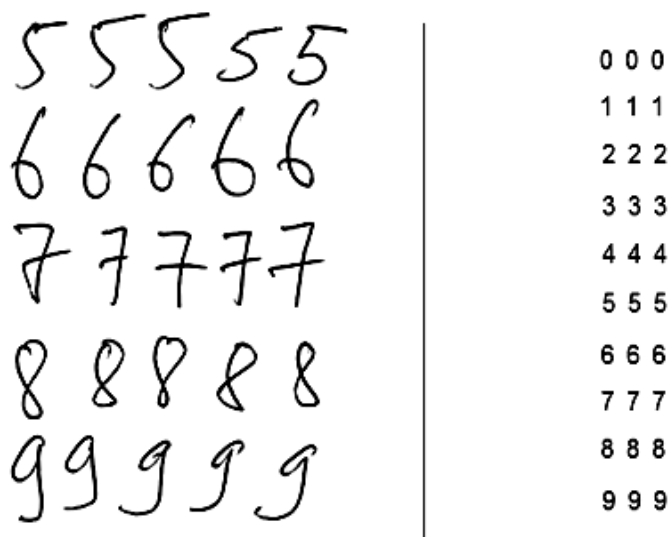
Hình 3.9 Ký tự số viết tay cho máy nhận dạng

Bảng: Kết quả đối sánh sử dụng độ đo khoảng cách Euclide.

Mẫu Ký tự	0	1	2	3	4	5	6	7	8	9
0	63.73	114.98	123.23	97.78	122.36	130.28	172.9	161.54	132.29	94.34
1	82.81	51.75	81.5	91.23	62.31	71.29	75.54	98.18	107.71	164.83
2	96.04	88.82	43.69	129.41	107.33	98.79	155.68	134.38	171.39	178.17
3	81.75	111.58	88.86	72.91	110.92	138.25	132.78	87.63	147.84	108.88
4	119.96	76.96	173.73	158.01	55.56	116.11	167.26	99.88	185.07	202.79
5	144.68	199.79	139.73	124.12	179.73	66.34	145.69	169.94	226.97	191.51
6	103.31	173.64	106.95	80.67	150.66	120.46	65.29	146.15	122.56	106.24
7	156.03	206.05	166.55	103.29	136.72	157.35	172.52	95.33	189.87	152.28
8	80.09	136.72	89.5	75.34	120.41	119.12	104.45	117.33	40.84	113.59
9	130.19	157.74	208.57	104.95	195.96	202.68	199.57	150.13	179.47	91.11

3.3.3. Nhận dạng bằng điểm

Cho máy học các ký tự số viết tay và in được người vào điểm viết mẫu



a) Các ký tự điểm viết tay

b) Các ký tự số in để nhận dạng mã số

Hình 3.11 Các ký tự cho máy học

Kết quả nhận dạng bảng điểm đạt khoảng 96%

Dưới đây là giao diện ứng dụng nhận dạng bảng điểm và kết quả

Training Recognition Application

Academic Transcript

SỞ GD & ĐT QUẢNG NGÃI
TRƯỜNG TIỂU HỌC LÊ ANH

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BẢNG ĐIỂM KIỂM TRA 1 TIẾT - NĂM HỌC 2011 - 2012

MÔN: VẬT LÝ NGÀY KT: 4-10-2011

GV CHẤM: TÔN THẤT ĐIỀU TỔ QUẢN VỤ: Phu

Results



STT	MSHS	HỌ VÀ TÊN	LỚP	ĐIỂM		GHI CHÚ
				SỐ	CHỮ	
1	11008	Võ Thị Thuý Anh	10A1	9	Chín	
2	11009	Nguyễn Thị Anh	10A1	9	Chín	
3	11010	Lê Quốc Bảo	10A1	6	Sáu	
4	11011	Nguyễn Chí Minh	10A1	8	Tám	
5	11012	Nguyễn Hữu Chính	10A1	7	Bảy	
6	11013	Ngô Hùng Phú Cường	10A1	7	Bảy	
7	11014	Nguyễn Văn Dung	10A1	9	Chín	
8	11015	Bùi Đình Chung	10A1	7	Bảy	
9	11016	Nguyễn Phước Duy	10A1	6	Sáu	
10	11017	Lê Thị Kim Chung	10A1	8	Tám	
11	11018	Đinh Nguyễn Minh Hiền	10A1	6	Sáu	
12	11019	Nguyễn Mạnh Cường	10A1	7	Bảy	
13	11020	Võ Thị Hồng Diệp	10A1	5	Năm	
14	11021	Lê Văn Huy	10A1	7	Bảy	
15	11022	Nguyễn Bá Huỳnh	10A1	7	Bảy	
16	11023	Hồ Quốc Kiên	10A1	5	Năm	
17	11024	Bùi Văn Lợi	10A1	7	Bảy	
18	11025	Trương Hoàng Ngọc	10A1	7	Bảy	
19	11026	Nguyễn Văn Phương	10A1	5	Năm	
20	11027	Trần Thị Phương	10A1	6	Sáu	
21	11028	Nguyễn Thị Phước	10A1	7	Bảy	
22	11029	Huỳnh Văn Quang	10A1	8	Tám	
23	11030	Nguyễn Thanh Siêng	10A1	6	Sáu	
24	11031	Nguyễn Phú Thành	10A1	7	Bảy	
25	11032	Trần Thị Thu	10A1	9	Chín	
26	11034	Nguyễn Công Tuấn	10A1	8	Tám	

Recognition Character

Hình 3.12 Giao diện nhận dạng bảng điểm

3.4. Đánh giá kết quả

Kết quả nhận dạng cho thấy phương pháp sử dụng đường cong B-Spline để nhận dạng ký tự in và ký tự viết tay là khá hiệu quả. Nhận dạng ký tự số in cùng kích thước đạt tỷ lệ chính xác 100%, nhận dạng ký tự viết tay đạt 98%. Đặc biệt độ chính xác càng tăng khi các mẫu có kích thước càng lớn vì độ lệch đặc trưng của các mẫu tỉ lệ thuận với kích thước, kích thước càng lớn độ lệch đặc trưng

càng cao và ngược lại. Tuy nhiên, trường hợp người dùng cho hệ thống nhận dạng các mẫu có kích thước bé, độ lệch đặc trưng giảm, dẫn đến việc hệ thống sẽ nhận dạng nhầm lẫn giữa các mẫu.

Kết quả nhận dạng của hai nhóm viết tay khác nhau đạt tỷ lệ thấp. Cho nhóm A viết mẫu và nhóm B viết nhận dạng, kết quả thu được là khoảng 50%.

Nhận dạng bảng điểm, kết quả đạt 96%. Độ chính xác khi nhận dạng bảng điểm không như mong muốn, nguyên nhân là do chất lượng ảnh của bảng điểm kém, các ký tự điểm viết tay bị trùng lên các đường biên do đó kết quả tiền xử lý hình ảnh không đạt hiệu quả cao dẫn đến máy ra quyết định nhận dạng điểm sai.

KẾT LUẬN

1. Kết quả đạt được

Qua một thời dài nghiên cứu kết hợp với học hỏi kinh nghiệm từ những người đã nghiên cứu trước về lĩnh vực này, luận văn đã hoàn thành và đạt được những kết quả sau:

Về lý thuyết: Luận văn đã trình bày tổng quan các vấn đề về đường cong tham số B-spline, xử lý ảnh dùng trong nhận dạng chữ viết tay, nêu được giải pháp kỹ thuật để xây dựng được hệ thống nhận dạng.

Về thực tiễn: Luận văn đã xây dựng được một hệ thống sử dụng đường cong B-Spline để nhận dạng các ký tự viết tay. Kết quả thực nghiệm cho thấy khả năng nhận dạng của phương pháp này là rất chính xác, đặc biệt các ký tự có kích thước càng lớn thì việc nhận dạng càng tốt. Cách tổ chức lưu trữ các mẫu chuẩn là rất quan trọng, sẽ ảnh hưởng rất lớn đến độ chính xác khi nhận dạng.

2. Hướng phát triển của luận văn

Bên cạnh những vấn đề đã đạt được như đã trình bày, đề tài cần được phát triển thêm các vấn đề sau:

- Hoàn thiện các chức năng của chương trình.
- Hoàn thiện chức năng nhận dạng bảng điểm để triển khai cho phép nhận dạng bảng điểm tự động ở các trường học.
- Phát triển ứng dụng để nhận dạng văn bản viết tay, đặc biệt là nhận dạng văn bản viết tay tiếng việt. Phát triển để có thể nhận dạng văn tay...